DOCUMENT RESUME

ED 035 454 PS 002 767

AUTHOR	Baldwin, A. L.; Frank, S. M.
TITLE	Syntactic Complexity in Mother-Child Interactions.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
PUB DATE	Mar 69
GRANT	OEG-6-10-326
NOTE	16p.: Paper presented at the meetings of the Society for Research in Child Development, Santa Monica, California, March 1969
EDRS PRICE	EDRS Price MF-$0.25 HC-$0.90
DESCRIPTORS	Comparative Analysis, *Complexity Level, Kernel Sentences, *Language Development, Parent Child Relationship, Socioeconomic Status, *Syntax, Transformation Generative Grammar

ABSTRACT

To find out what factors are involved in a child's learning of syntax, investigators studied the syntactic complexity of the language a mother and child use when talking to each other. The complexity measure used was one developed by Dr. Sheldon Frank and Dr. Harry Osser, and is based on the concepts of generative grammar and transformations. Language samples were collected from mothers alone in an interview and mothers and children together in a play session. There were two groups of mother-child pairs: one from Harlem (black lower class) and one from Washington Square (white middle class). Analysis of the language samples indicated that all the mothers greatly reduced their syntactic complexity when talking to their children, but each mother's language was still more complex than that of her child. There was no difference between the Harlem mothers and the Washington Square mothers in their syntactic complexity in the interview, but the Harlem mother-child interactions were less complex and more didactic than those of the Washington Square group. Finally, the Harlem children seemed to articulate less clearly than the Washington Square children and their mothers had more difficulty in understanding them. There were more requests for clarification in the Harlem interactions. (MH)

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED035454

Syntactic complexity in mother--child interactions

A. L. Baldwin and S. M. Frank

You have just heard about our procedure for describing some aspects of the interaction of the mother and the child and some of the features of the interactions we have observed. These descriptions have been concerned with the psychological impact of the mother and child's verbal interaction, but not with the actual language used. We have been interested mainly in what the mother and child communicate to each other, not with the precise way they say it. Language development is obviously one aspect of cognitive development, and therefore we have been interested in the language the mother and the child use in their interactions because it almost surely has an influence on the development of the child's language skills.

We have studied the level of syntactic complexity of the language on the assumption that such an index may give us useful information about the factors involved in the child's learning of syntax. We have been particularly interested, therefore, in the level of syntactic complexity used by the mother in interacting with the child and the complexity of his language in interacting with her. When we have developmental data, we hope to find that the child's growth in syntactic complexity is related to the difference between his language usage and that of his mother.

PS002687

[1] This research has been supported by the Office of Education grant No. OE-6-10-326. This paper was presented to the meetings of Society for Research in Child Development in Santa Monica March 1969

The measure of complexity that we have used is one developed by Dr. Sheldon Frank and Dr. Harry Osser in their study of language development in Negro children in Baltimore. I feel embarassed to be reporting on this measure because I am not a linguist, and I will not be able to defend this measure adequately or even answer many questions about it's details. We chose this method (1) because it is applicable in these naturalistic situations, (2) because it is a differentiated measure that makes relatively fine discriminations among sentences, (3) because it depends on many elements in the sentence so should reflect many potential aspects of language development, (4) because it seems to be closely related to our intuitive feelings about the complexity of a sentence, and (5) finally because Dr. Frank was available as a consultant to our project. He trained the coders, supervised the scoring of the records, and has collaborated fully with us in the analysis of the results.

Now let me try to describe the measure itself. It is based on the concepts of generative grammar although like many psycholinguistic studies the syntactic theory on which it is based has become somewhat outmoded by the time the study is done.

Chomsky as of 1957 described sentences as kernel sentences or as more complex ones derived from kernel sentences by transformations. The measure Dr. Frank developed is based upon the idea that the kernel sentence consists of a nucleus, the bare bones of the noun phrase and verb phrase, with perhaps the addition

of one or more optional elements like an adverbial phrase. Then the kernel sentence is transformed into the final output through the application of one or more transformations. Each transformation involves the following elements: addition of an element, transposition of an element from one place in the sentence to another, the deletion of some element, or the addition of supersegmental elements like an intonational change. The basic assumption for the scoring is that each of these changes in a sentence, whether in the kernel or through transformations, is of equal difficulty. Therefore the score is the number of such additions, deletions, transpositions, and intonational changes.

Let me give you a few examples:

A complete nucleus like "we played" receives a score of 2, one for the noun phrase, one for the verb phrase. Some sentences that occur in conversation may lack one of these elements, e.g. "Huh"; or "a car" in answer to the question "What is that". These utterances receive scores of 1 and there are many in our records.

"I never played only one time" is scored 5. The nucleus "I played" receives 2, the additional optional elements are "never", "one time", and "only". Each adds 1 to the score.

"We sing songs and play" is scored 4; 2 for the nucleus and 2 for the transformation that adds a second sentence to the first with an "and". If the second sentence had been complete, "we sing songs and we play", the score would have been 5. One might argue that the original form is obtained from the latter

by a deletion of the noun phrase from the second sentence and therefore ought to receive an additional score. Here Frank departs from the strict logic of transformational grammar and argues that the many deletions that occur in generative grammar do not all count as additional complexity. This is certainly a debatable question, but the important thing for the present study is that the scorer need not debate it. Each transformation has an assigned score so that all instances of it are scored the same way.

I'll give you some more examples of sentences and their scores without analyzing each one.

We sit on the circle and stand on the circle.	Score 6
What is that?	Score 5
You set them on the floor because they can't stand on the rug.	Score 12
Well, you go over and see what it is.	Score 16
Ain't that the thing that say, "oink, oink, oink."	Score 11
Yea, stand it up and see if you know what it is.	Score 21

The analysis can be made in various ways to study the appearance of particular transformations and specific optional elements, but in this study we have concentrated on a quantitative score, the mean score per sentence, which will be referred to as mean complexity, and the total number of different transformations used in the sample of sentences analyzed.

Any analysis of the features of a naturalistic situation runs into problems of sampling. The purely quantitative side of sampling is relatively easy to deal with. We began with the analysis of a sample of fifty consecutive sentences beginning with the start of the second third of the record. The correlations for mean complexity between the fifty sentence sample and the total sentences is shown in Table 1 of the handout. It is above .70 for three of the four samples of people, Harlem mothers, Harlem children and Washington Square children, but for Washington Square mothers the correlation was only .40. We then increased the sample to 100 sentences. The correlation of this larger sample with the total is above .80 for every sample.

The qualitative aspect of sampling is more difficult. Suppose one mother-child pair play exclusively with the jig-saw puzzle, another plays with the doll family. If the fantasy connected with the doll play tends to produce more complicated sentences than the conversation about the pictures on the jig-saw puzzle, the difference between the two pairs might not reflect any basic difference in complexity of speech, but merely the objects in the room that they played with.

To investigate this problem we categorized each sentence in two ways. First, what toy was the child playing with when it was said, and second, what object was the subject matter of the sentence. There were several residual categories: when the sentence was about some external object unrelated to the play room; when it was about two things at once; when it concerned the

PS 002767

behavior of the people involved. There were occasional uncodable sentences.

Since some mother-child pairs did not play with every toy in the room. there are some categories that are absent in each record. These missing values were estimated from the mean of that pair's score, and the mean score for that object of play. Such estimates would not enhance the difference among the various object categories and contributed nothing to the error variance. The error variance was then estimated with a degree of freedom that was reduced by the number of missing items. This procedure is a debatable one, but it was the only one we could devise and it should roughly indicate whether the problem of sampling is serious. Table 2 shows the analysis of variance for the four samples. As you can see, the effects of the toy being talked about on the complexity of the sentence is significant, but not always very big. Since it is significant, we calculated a revised score for each person in the sample that is an estimate of what the average complexity for that person would be if each category of sentence was uttered with equal frequency. In other words, it makes the content of the utterances of different people comparable in terms of what they talked about. When we look at the results we will see how this revised measure affects the interpretation.

Results

Table 3 shows the comparison of the mother's language during the interview and during the play session with the child.

The most striking feature is the tremendous reduction in syntactic complexity when the mother is playing with the child. It is obvious that every mother in the sample modulates the grammatical complexity of her language when talking to her child. The average complexity of the sentences during the interview was about 13, comparable to the example given earlier, "You set them up on the floor because they can't stand on the rug." With the child the complexity of the mother's language is about 5. Furthermore, there is a correlation of .45 between the complexity in the two situations. With an N of 12, this correlation is not significant.

The next important finding is that the mother's syntactic complexity during the play sessions is consistently higher than the child's. Mothers clearly modify their language with children, but they do not talk as simply as the child does. The implications of this fact are not clear. Possibly this difference between the mother's and the child's complexity of grammar is a source of some of the child's language learning. How big a difference is optimal is unknown. One would suspect that too little difference would not be sufficiently stimulating, and that too big a difference would merely leave the child not understanding what his mother is saying.

Table 4 shows the same comparisons for the number of different transformations used. Here the size of sample is critical because obviously there will tend to be more different transformations used in a large sample of sentences than in a small sample. Therefore the comparison of the interaction with

the interviewer is based on a fifty sentence sample and the comparison of mother and child on a total sample. All the previous findings on complexity are confirmed in this table.

These then are the main overall findings about syntactic complexity in mother-child interactions. These findings hold true for the Harlem sample and the Washington Square sample. There are some differences between the two samples, generally much smaller than those we have been talking about, but worth reporting.

First, in the interview situation, there is no difference between the two samples. If Harlem residents speak in a restricted code by comparison with middle class white people, they don't show it in this situation.

Robert Hess and his co-workers found clear differences between different socio-economic levels in the complexity of the mother's language in interacting with the adult experimenter. His measure of complexity is not the same as Frank's, but measures many of the same elements and would surely be highly correlated with it. Since the Harlem sample we have been studying comes from varying social classes from ADC to Ph.D, the equal language complexity of the two samples in the interview situation does not directly contrast to Hess's findings. But since there is certainly a difference between the average social class status of our two samples, the lack of agreement with Hess's findings raises a problem.

We do find a difference, however, between the two samples in the play session. Both the mother and the child in the Washing-

ton Square sample use a slightly but significantly greater complexity of grammar in their interaction. The difference is about .5 units. The difference between the mother and the child, however, is almost identical in the two samples. Mothers speak with 1.9 units more of complexity than their children. There is another striking difference in the two samples in the correlation between the mother and the child's complexity of language. For the Harlem sample this correlation is +.72, significant at 5% level, while for the Washington Square sample the correlation is zero.

To summarize the empirical findings, we find large differences between the mother's language in the interview and in the play session, and large differences between the mother's and child's language in the play session. These hold for all mother-child pairs. There are much smaller sample differences with the Washington Square sample, both mother and child speaking slightly more complex language than the Harlem sample, but with zero correlation between mother and child. These differences are still obscure.

Discussion:

Instead of trying to interpret the language complexity separately, I would like to try a broader interpretation of these sample differences. From our observations of these mother-child interactions, we have the impression that one difference between the two samples is an emphasis by the Harlem mothers on teaching the child specific skills, particularly

school related skills like the names of letters, labels for animals etc; while the interactions of the Washington Square sample are more casual and spontaneous. There are a number of measured characteristics of the interactions that support this interpretation: (1), the greater percentage of questions in the Harlem sample, more of them from the mother than the child; (2), the greater frequency of utterances involving permanent information in the Harlem sample, but more fantasy information in the Washington Square sample.

If there is a greater frequency of direct didactic teaching, we might also expect some of the differences in syntactic complexity that were obtained. It seems likely that such teaching is in the form of simple, clearly understood questions and statements, whereas the language in fantasy play may well tend to be more complicated. Also, if the mother is directly instructing the child, it could well be that her language is closely attuned in complexity to that of the child. Thus, the hypothesis of more didactic teaching in the Harlem sample seems consistent with the slightly lower complexity of the Harlem interactions and the high correlation between mother and child complexity in the Harlem interactions.

It is easy to see several reasons for didactic teaching in the Harlem sample. The Harlem mothers that we observed were deeply concerned about the child's language development and preparation for school. Some had older children who were having trouble in school. They did not feel the confidence that

the Washington Square mothers did that the child would develop adequately without special effort. Also, they were participating in a study of language development with white experimenters and may well have been motivated to focus the interactions more on language development because that's what we told them we were interested in. The Washington Square mothers can probably take university professors or leave them alone and thus be more casual in the play session.

A second impression is that the Harlem children articulate less clearly than the Washington Square children. The basic evidence is that there are significantly more uncodable utterances from the Harlem sample. Several possible artifacts may be producing this apparent difference. (1) We were constantly striving to improve our audio equipment and the Harlem sample was seen six months earlier in real time than the Washington Square sample. (2) The transcriber was white and probably had trouble with the Negro dialect, particularly at spots where the tape was noisy. On the other hand, we had a Negro college student from the city go over all the Harlem tapes to fill in words and phrases that he could understand. Still there are more unclear utterances. We have the impression, moreover, that the Harlem mothers themselves did not understand their children's speech as often as did the Washington Square mothers. We can analyze the records for the number of untranscribable blanks followed by a mother's request for clarification, but have not yet done so. We do know, however, that there are significantly

more requests for clarification and clarifying statements in the Harlem interactions. Since these questions and their answers are simple sentences, their greater frequency could also contribute to the lower syntactic complexity in the Harlem interactions.

In general, we feel that there are real, understandable differences between the two samples, but we are much more impressed with the homogeneities that describe the general structure of verbal interaction between mother and child in a play session. There are norms of interaction, for example response-demanding utterances by the high power person are more frequently responded to than those by the low power person in the interaction (see data on utterances ignored). Questions tend to demand statements in response, generally of the same content; e.g. requests for clarification elicit clarifications. If these general structural rules can be made explicit, it may even be possible to develop a crude mathematical model of mother-child interactions in which there are many fewer independent parameters than there are categories. These parameters are what will best describe the differences between different mother-child pairs. This is the exciting vista of research that we glimpse after the analysis of these interactions.

Baldwin and Frank

Syntactic Complexity in Mother-Child Interactions

Table 1

Effect of Sampling as Measured by Syntactic Complexity

	odd even	50 sen.	r with Total	100 sen.	r with Total	Total
Harlem						
Mother		4.36	.78	4.38	.81	4.55
Child	.87	2.65	.89	2.63	.95	2.66
Wash.Sq.						
Mother		5.19	.40	5.03	.89	5.06
Child		3.24	.71	3.19	.97	3.21

Syntactic Complexity in Mother-Child Interactions

Table 2

Relation of syntactic complexity to play object

1. Object being talked about

	Harlem		Washington Square	
	Mother mean sq.	Children	Mother	Children
Subject	6.46	7.72	2.22	4.86
Object	6.21^{xxx}	$.95^{x}$	6.30^{xxx}	1.42^{xx}
Error	.67	.42	1.27	.52

2. Object being played with

	mean sq.			
Subject	6.62	7.66	2.39	5.05
Object	6.10^{xxx}	1.02^{xx}	6.50^{xxx}	$.98^{x}$
Error	.75	.31	.58	.49

x Obj.mean sq./ error is significant at 5% level

xx mean sq/ error is significant at 2% le

xxx mean sq./error is significant at 1% le

Syntactic Complexity in Mother-Child Interactions

Table 3

Average complexity of Mother and Child.

	Harlem		Washington Square
Mother (Interview)	13.22	=	13.43
	r = .44		r = .46
Mother (Interaction with Child)	4.55	<	5.06
	r = .77		r = .-02
Child (Interaction with Mother)	2.66	<	3.21

> sign at .05 level >>> sign at .001 level =T>

Syntactic Complexity in Mother-Child Interactions

Table 4

Number of Transformations used by Mother and Child

	Harlem		Washington Square
Mother (interview) 50 sentences	35.58	=	36.08
	⩛		
Mother (Interaction with child) 50 sentences	18.33		23.50
Mother (Interaction with child) Total sentences	39.00	n.sign	43.67
	⩛		⩛
	r = .20		r = -.36
Child (Interaction with mother) Total	19.00	<	25.25